

4. Robótica de Servicio

4.1. El espacio de la Robótica

La robótica estudia los mecanismos físicos controlados por un programa de cómputo que interactúan con los seres humanos y/o el entorno y realizan tareas que involucran un proceso informacional y una acción motora o perceptual de manera coordinada. Aunque la palabra “robot” proviene de la ciencia ficción y tiene como antecedentes los autómatas mecánicos que se han construido a lo largo de la historia,¹ la robótica como disciplina científica y tecnológica se inicia con la invención de los brazos autónomos para asistir la manufactura en líneas de producción, especialmente en la industria automotriz, con el robot *Unimate* en 1961.² Esta invención dio lugar al desarrollo de una industria de grandes dimensiones con gran impacto económico y estratégico en el mundo.

A partir de estos robots iniciales, en conjunto con el desarrollo de las disciplinas relacionadas, como el Procesamiento de Señales, la Teoría de Control y la Inteligencia Artificial, durante las últimas décadas se han diseñado y construido una gran cantidad de robots o dispositivos robóticos que se pueden conceptualizar o clasificar en relación a dos dimensiones principales: i) autonomía y ii) complejidad del entorno. La primera depende de la riqueza

¹ <https://en.wikipedia.org/wiki/Robotics>

² <https://en.wikipedia.org/wiki/Unimate>

de sus sensores y de la variedad y estructura de sus mecanismos motores e informacionales o “mentales” y, consecuentemente, de la riqueza de estímulos o informaciones que el robot es capaz de percibir y/o interpretar, y de la cantidad y estructura de los tipos de acciones que puede realizar para conseguir sus objetivos de manera autónoma.

La segunda corresponde a la variedad y variabilidad del entorno que el robot es capaz de enfrentar e incluso transformar. Esta dimensión no se trata de la extensión espacial donde los robots pueden funcionar e interactuar, sino de la cantidad y complejidad de entornos cualitativamente diferentes; por ejemplo, aunque las hormigas, las cucarachas o los osos de agua son ubicuos en el planeta, sus micro-entornos locales, en relación a su escala, son muy similares en todos lados, en oposición a los entornos globales complejos, cualitativamente diferentes, como los que podemos enfrentar los seres humanos. Aunque la diversidad y la variabilidad del entorno son dimensiones diferentes están relacionadas ya que entornos simples tienden a variar poco y entornos complejos mucho, por lo que se pueden colapsar en una sola dimensión a la que nos referimos como “complejidad”. Mientras mayor sea la autonomía del robot al actuar e interactuar en entornos complejos, mayor será su inteligencia.

De acuerdo con estas consideraciones es posible describir el mapa de la robótica en el espacio definido por estas dos dimensiones, como se ilustra en la Figura 4.1. En el origen están los robots con un solo tipo de acción en un entorno fijo, como los brazos industriales y en la esquina superior derecha los robots humanoides de la ciencia ficción. Asimismo la gran mayoría de los robots actuales y potenciales estarían sobre una nube a lo largo de la diagonal principal ya que la complejidad del robot debe ser coherente con la complejidad de su entorno, aunque habría muchas excepciones, como los robots completamente autónomos dirigidos a actuar en entornos de complejidad baja o moderada.

Por otra parte, desde una perspectiva funcional y de su utilidad, el espacio de los robots se puede dividir en tres grandes regiones o semiplanos paralelos verticales de izquierda a derecha: i) industriales, ii) controlados de manera remota y iii) completamente autónomos. La primera región corresponde a los robots de las líneas de producción cuyo entorno no varía o varía muy poco y su protocolo de acción está definido de antemano, por lo que sus procesos de percepción, toma de decisiones y acción intencional son muy limitados. La segunda incluye mecanismos con la capacidad de explorar y manipular el entorno pero la información que obtienen a través de sus sensores se envía a un operador humano quien los controla de manera remota y decide sus acciones. Estos robots tienen grandes aplicaciones actuales y potenciales, ya que pueden operar en ambientes peligrosos o inaccesibles para los seres humanos, como inspeccionar el núcleo de un reactor nuclear o buscar sobrevivientes después de un terremoto. La tercera categoría corresponde a robots capaces de percibir e interpretar el mundo y tomar decisiones de manera autónoma, principalmente en entornos complejos, aunque, salvo algunas excepciones, como los vehículos autónomos, los robots de esta última categoría están todavía en los laboratorios de investigación. En la Figura 4.1 se ilustran algunos ejemplos de robots en este espacio.



Figura 4.1. Ejemplos de robots en el Espacio de la Robótica.

4.2. Robots de Servicio

Con base en el trabajo de investigación y desarrollo de la comunidad mexicana de computación, en este capítulo nos enfocamos a los robots de servicio, un tipo de robot de media o alta autonomía capaz de asistir a los seres humanos en entornos de complejidad baja o moderada. Es importante enfatizar que estos robots son diferentes a los industriales ya que su objetivo es simplificar el trabajo humano en casas, oficinas, hospitales, etc., para lo cual tienen que moverse de manera intencional, evitando obstáculos,³ reconocer e identificar a los humanos, comunicarse a través del lenguaje hablado, ver, tomar, llevar y entregar objetos entre otros tipos de acciones o conductas intencionales. Los robots de servicio deben también ser capaces de hacer diagnósticos, tomar decisiones y hacer planes para llevar a cabo sus funciones de manera exitosa. Se espera que en un futuro cercano dentro de la casa habrá robots pequeños que limpien, aspiren y trapeen el piso; robots fijos encargados del lavado y planchado de la ropa; y robots de tipo humanoide que fungirán como asistentes generales. En el exterior habrá robots que cortarán el pasto y otros que harán rondines de vigilancia. Así como se incorporaron a la vida cotidiana los televisores, las computadoras y los teléfonos celulares, los robots de servicio también lo harán y llegarán a ser muy familiares, si se cumplen las expectativas que tenemos actualmente. Se espera que una de las aplicaciones principales de esta tecnología en el mediano plazo sea asistir a adultos mayores como acompañantes, para explicar acciones rutinarias, llevar la lista y toma de medicinas, apoyar en el ejercicio físico, proveer información de calendario y hora, proveer información espacial, detectar accidentes, como caídas, o mejor aún, mantener la casa segura para evitar accidentes o situaciones de riesgo (ej. apagar la estufa, recoger objetos del suelo, etc.). En el ámbito hospitalario para transportar medicinas, análisis y alimentos, guiar a los médicos en las visitas y monitorear los signos vitales de los pacientes.

³ En la terminología de la especialidad esta acción se conoce como “navegar” y a la sub-disciplina que aborda la problemática relacionada como “navegación”.

4.3. Conceptualización y arquitectura de Robots de Servicio

En el desarrollo de los robots de servicio se han observado de forma genérica tres paradigmas o enfoques de robots móviles, cada uno de los cuales determina hasta cierto punto la arquitectura del robot. El primero, llamado “deliberativo”, se desarrolló desde los años sesenta del siglo pasado, basado en el paradigma simbólico y representacional del programa de la Inteligencia Artificial. Este enfoque parte de una representación del mundo y a través de un proceso de planeación se determina la secuencia de acciones que el robot debe realizar para lograr sus objetivos. Un ejemplo de este tipo de arquitecturas es la del robot Shakey⁴ desarrollado en la Universidad de Stanford. El segundo enfoque, llamado “reactivo”,⁵ se desarrolló en los años ochenta en el Instituto Tecnológico de Massachusetts (MIT) por el investigador Rodney Brooks; en éste no se cuenta con una representación del medio ambiente, ni tampoco con módulos de planeación de acciones ni movimiento, sino que se utilizan varios módulos que responden a estímulos específicos provenientes del entorno y se generan salidas inmediatas las cuales son combinadas o seleccionadas por un árbitro para generar lo que el robot debe hacer bajo ciertas circunstancias. Los módulos pueden utilizar máquinas de estados, redes neuronales o campos potenciales. El tercer enfoque, llamado “probabilista” se basa en el concepto de que tanto el sensado como las acciones del robot en ambientes complejos son inciertos y se representan mediante variables aleatorias, las cuales pueden ser manipuladas para obtener un plan de acción robusto. Un ejemplo de este enfoque es el automóvil autónomo que obtuvo el primer lugar en el *DARPA Grand Challenge* del 2005 desarrollado por Sebastian Thrun de la Universidad de Stanford.⁶

⁴ <http://www.ai.sri.com/shakey/>

⁵ https://en.wikipedia.org/wiki/Behavior-based_robotics

⁶ Thrun, S., et al. (2006). **Stanley: The Robot that Won the DARPA Grand Challenge.** *Journal of Field Robotics*, 23(9): 661–692.

4.4. Modelo conceptual y niveles de sistema en Robótica

Desde el punto de vista de su uso, diseño y construcción, los robots se pueden conceptualizar en tres niveles de sistema: i) funcional, ii) de dispositivos y algoritmos y iii) de implementación.⁷ El primer nivel corresponde a la perspectiva de los seres humanos o usuarios que interactúan con el robot: qué servicios ofrece en términos prácticos, cómo se estructuran las tareas que puede realizar y cuál es la estructura de la comunicación o la interacción con los agentes en su entorno. Si las expectativas que tenemos para los robots autónomos se llegaran a materializar éstos se comprarán en agencias, como las de coches, y vendrán con un manual de usuario con la especificación de sus conductas básicas y las formas de combinarlas para realizar tareas compuestas. Dicho manual correspondería con la especificación del nivel funcional. Este nivel funcional tiene también un impacto directo en la arquitectura del robot ya que determina en buena medida cómo se utilizan los dispositivos y algoritmos robóticos así como la forma en que el robot interactúa con otros agentes y cómo responde a los eventos que ocurren en su entorno, tanto naturales como los producidos por la acción de otros agentes de manera intencional.

El nivel de dispositivos y algoritmos corresponde al diseño y construcción de artefactos robóticos junto con los algoritmos o programas que los habilitan: brazos con manos capaces de tomar objetos frágiles con la fuerza suficiente para sostenerlos pero sin romperlos; cabezas con cámaras y algoritmos de visión para reconocer objetos y determinar su posición y orientación, que en conjunto con los brazos y las manos habiliten al robot para tomar objetos; plataformas móviles o piernas que le permitan al robot trasladarse de una ubicación a otra, asociados a algoritmos que le permitan construir un mapa de su entorno y moverse evadiendo obstáculos; dispositivos y algoritmos que lo habiliten para interactuar o comunicarse a través del lenguaje hablado, o reconocer ruidos y voltear hacia la fuente sonora y ponerla en su

⁷ Pineda, L. A., Rodríguez, A., Fuentes, G., Rascon, C., Meza, I.V. (2015). **Concept and Functional Structure of a Service Robot**. *Int J Adv Robot Syst*, 12:6. doi: 10.5772/60026

foco de atención. En este nivel también se incluyen sistemas de representación del conocimiento que le permiten hacer inferencias conceptuales y sistemas de razonamiento deliberativo para hacer diagnósticos, tomar decisiones y planear sus acciones futuras. Este nivel de sistema es el que se asocia más directamente con la práctica de la robótica, o de manera más coloquial, con lo que hacen los científicos y tecnólogos “roboticistas” o los “roboteros”. Asimismo, muchos dispositivos y algoritmos en este nivel se apoyan en técnicas de análisis de señales, reconocimiento de patrones, aprendizaje de máquina y teoría de control, entre otras, por lo que desde la perspectiva de la robótica estas disciplinas se subordinan al nivel algorítmico.

Finalmente el nivel de implementación corresponde a los programas de apoyo que se requieren para habilitar los diversos dispositivos y funciones, pero que no son tecnología robótica propiamente. Se pueden mencionar aquí los sistemas operativos y sistemas de comunicaciones que se utilizan para integrar las partes físicas o computacionales del robot como un todo. Esta tarea la realizan programadores de sistemas como apoyo a los roboticistas.

Todos los robots se pueden conceptualizar en términos de estos tres niveles de sistema pero es frecuente que sólo el nivel de dispositivos y algoritmos se considere explícitamente, en particular en el diseño y construcción de robots simples. Sin embargo, cuando el robot está constituido por varios dispositivos y requiere de varios algoritmos complejos, que tienen que operar de manera coordinada, es indispensable distinguir los tres niveles de manera explícita y abordarlos con metodologías específicas y grupos de investigación y desarrollo enfocados a los niveles funcional y de implementación, de manera adicional a los grupos enfocados a cada una de las especialidades en el nivel de dispositivos y algoritmos.

4.5. Aproximaciones al nivel funcional en México

En esta sección se presentan tres métodos para organizar y estructurar robots de servicio en el nivel funcional; estos enfoques son los adoptados y desarro-

llados por los grupos mexicanos que trabajan con robots de servicio, como sigue: i) máquinas de estado, ii) lenguajes de especificación de tareas y iii) procesos de decisión de Markov. El primero se ha adoptado por el grupo Pumas de la Facultad de Ingeniería de la UNAM y se implementa en el robot Justina; el segundo se ha desarrollado por el Grupo Golem⁸ del Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas (IIMAS) de la UNAM y se ha implementado en los robots Golem, Golem-II+ y Golem-III, y el tercero se ha adoptado y desarrollado en el Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE) y se ha implementado en los robots Markovito y Sabina.

4.5.1. Máquinas de Estados Finitos

La forma más directa para programar la conducta de un robot de servicio es mediante el uso de máquinas de estados. En éstas se tiene una representación del estado en el cual se encuentra el sistema y dadas ciertas condiciones de entradas y el estado presente se calcula el estado siguiente. La parte fundamental de las máquinas de estado es el Autómata de Estados Finitos (AEF) que especifica y guía la conducta. Por ejemplo, la conducta de un robot omnidireccional con dos motores que le permiten ir hacia adelante, hacia atrás y hacer giros de 45 grados hacia la derecha e izquierda, que cuente con dos sensores que le permiten detectar obstáculos (izquierdo y derecho) y capaz de realizar un conjunto finito de acciones, se puede especificar como sigue: si el robot no detecta ningún obstáculo avanza en la dirección en que está orientado; si detecta un obstáculo con su sensor izquierdo se mueve hacia atrás y gira hacia la derecha 45 grados; si detecta un obstáculo con su sensor derecho se mueve hacia atrás y gira hacia la izquierda 45 grados; si detecta un obstáculo con los dos sensores entonces se mueve hacia atrás, gira dos veces hacia la izquierda, avanza hacia adelante y finalmente gira hacia la derecha dos veces. Las entradas y salidas del autómata se pueden asociar a diferentes dispositivos

⁸ <http://golem.iimas.unam.mx/>

y algoritmos, como redes neuronales implementadas en hardware (FPGAs)⁹ para interpretar una señal de entrada o a campos de potenciales, para decidir una acción de navegación que se ejecuta cuando el control de estados se mueve de un estado al siguiente. Estas conductas pueden ser externas, como moverse hacia atrás o girar, o internas, como consultar a una base de datos de conocimiento. En el caso de que la interpretación y la conducta sean concretas el robot será “reactivo” y en la medida en que sus conductas internas sean más ricas tendrá una orientación más representacional y deliberativa. El robot Justina, el cual se ilustra en la Figura 4.2 (a), es una realización de este modelo conceptual en el nivel funcional.



(a)



(b)



(c)

Figura 4.2. Robots de servicio desarrollados en México: (a) Justina (FI-UNAM), (b) Golem-III (IIMAS-UNAM), (c) Sabina (INAOE).

⁹ Savage, J., Cruz, J., Matamoros, M., Rosenblueth, D.A., Muñoz, S., Negrete, M (2016). **Configurable Mobile Robot Behaviors Implemented on FPGA Based Architectures.** En *16th IEEE Int. Conf. on Autonomous Robot Systems and Competitions*, pp. 317-322.

4.5.2. Lenguajes de especificación e interpretación de tareas

El modelo basado en autómatas de estado finito es limitado cuando la complejidad de la tarea y la comunicación aumenta, por lo que es útil incrementar el nivel de abstracción para especificar y ejecutar tareas complejas. Esto se puede lograr substituyendo la noción de “estado” por la noción más estructurada de “situación”. La diferencia es que mientras un estado tiene un contenido nulo de información, excepto por las etiquetas asociadas a sus transiciones y sus siguientes estados, una situación es rica informacionalmente, como en los cuadros de historietas en las caricaturas, donde hay uno o más personajes en el entorno o contexto. Una situación se puede caracterizar computacionalmente mediante la especificación de las expectativas de quien juega el papel del y_0 , en este caso el robot, las cuales corresponden con sus creencias acerca de los eventos naturales que pueden ocurrir en el mundo y a las creencias e intenciones que adscribe a sus interlocutores. La situación contiene también las acciones intencionales que el robot debe realizar cuando una de las expectativas se cumple. En este modelo la acción del robot se concibe como un tránsito continuo entre situaciones, como cuando se lee un libro de historietas, que corresponde con las tareas que el robot realiza en el mundo.

Por ejemplo, la conducta de un mesero se puede caracterizar de manera simplificada por un conjunto de situaciones, como estar en la puerta del restaurante esperando al cliente; llevar al cliente a su mesa; darle la carta y tomarle la orden; servirle; vigilar que todo esté en orden; llevarle la cuenta; cobrarle, llevarle el cambio y despedirle. En cada situación el mesero tiene un conjunto, de expectativas normalmente pequeño, así como la capacidad de realizar de manera inmediata la acción que se requiere cuando una de éstas se cumple. La noción de situación es una abstracción con una extensión temporal y espacial relativamente amplia y de ahí su utilidad: el mesero puede estar toda una hora moviéndose nerviosamente en el vestíbulo del restaurant en la situación “esperando al cliente” y sólo pasará a la situación de “llevando cliente a su mesa” cuando llegue un cliente. De hecho el agente cambia de situación justamente

cuando cambian sus expectativas. Por supuesto, hay situaciones complejas como “ver si todo está en orden” pero éstas se pueden partir a su vez en situaciones más simples, hasta llegar a las situaciones en que las expectativas correspondan con actos de interpretación concretos y las acciones se realicen directamente, como moverse o proferir una orden. Por estas razones, tareas relativamente complejas, como la conducta de un mesero, de un asistente en un supermercado o de un recepcionista en un hotel, se pueden modelar como una gráfica prototípica de situaciones, que a su vez se extiende o “desdobra” como una gráfica concreta cada vez que el robot lleva a cabo una tarea.

En este modelo las expectativas y las acciones guían a los dispositivos percepción y acción; sin embargo, las acciones pueden ser también internas o “mentales” y pueden consistir en hacer inferencias conceptuales o invocar a mecanismos deliberativos, como los sistemas de diagnóstico, toma de decisiones y planeación, que se utilizan de manera oportunista según los requerimientos del ciclo de la comunicación.

La pregunta obligada en este punto es qué pasa si no se cumple ninguna expectativa y/o sucede un evento inesperado, que además sea relevante. En este punto el robot pasa de estar situado a no estarlo y no sabe qué hacer. No estar situado es también perder momentáneamente la capacidad de interpretar y actuar en relación a la tarea o perder el contexto. Por lo mismo, cuando no se cumple ninguna expectativa u ocurre un evento inesperado se activa un modelo de diálogo cuyo propósito es simplemente volver a situarse, o alternativamente, invocar un ciclo de razonamiento para saber qué pasó y actuar de manera coherente para volverse a situar y realizar sus objetivos. Un robot que adopta este modelo conceptual es un robot situado.

Para implementar este modelo conceptual se pueden emplear lenguajes de programación orientados a la definición de tareas. Por ejemplo, el Grupo Golem desarrolló el lenguaje de programación de tareas robóticas SitLog.¹⁰

¹⁰ Pineda, L. A., Salinas, L., Meza, I.V., Rascon, C., Fuentes, G. (2013). **SitLog: A Programming Language for Service Robot Tasks**. *Int J Adv Robot Syst*, 10:358. doi: 10.5772/56906

Este lenguaje ofrece dos tipos de datos abstractos para especificar e interpretar tareas robóticas: el tipo *situación* y el tipo *modelo de diálogo*. El primero permite declarar directamente las expectativas y acciones de una situación y el segundo especificar una gráfica de situaciones que corresponde con la estructura prototípica de la tarea y, de manera paralela, a la estructura de la comunicación entre el robot y otros agentes en el entorno. El intérprete de SitLog consiste del intérprete de una red de transición recursiva (RTR),¹¹ que extiende a los AEF con una estructura de pila o *stack*, y del intérprete de un lenguaje funcional para la especificación e interpretación de expectativas y acciones, donde ambos intérpretes funcionan de manera coordinada. El intérprete de SitLog es el componente central de la arquitectura IOCA¹² utilizada en la serie de robots Golem y, en particular, en el robot Golem-III, el cual se ilustra en la Figura 4.2 (b). El intérprete de SitLog utiliza a su vez servicios de una base de conocimiento para llevar a cabo inferencias conceptuales¹³ para explorar la base de datos de conocimiento, así como inferencias deliberativas de manera oportunista.¹⁴

4.5.3. Procesos de Decisión de Markov

A medida que los robots de servicio interactúan en ambientes más variables y complejos se tienen que enfrentar a la incertidumbre tanto en su percepción del ambiente como en el resultado de sus acciones. Una forma de considerar y lidiar con dicha incertidumbre es a través de los procesos de decisión de Markov (MDPs, por sus siglas en inglés). Un MDP se puede ver como una extensión de las máquinas de estado que considera que las transiciones de un estado a otro, dada cierta acción, son probabilistas. Por ejemplo, si el robot realiza cierto movimiento como desplazarse 10 metros en línea recta, lo más

¹¹ https://en.wikipedia.org/wiki/Recursive_transition_network

¹² Pineda, L. A., Meza, I., Avilés, H., Gershenson, C., Rascón, C., Alvarado, M. and Salinas, L. (2011). **IOCA: Interaction-Oriented Cognitive Architecture**. *Research in Computing Science*, 54: 273–284.

¹³ Ver video en: <http://golem.iimas.unam.mx/lightkb/>

¹⁴ Ver video en: http://golem.iimas.unam.mx/opportunistic_inference

probable es que no avance exactamente 10 metros dado que se puede patinar, haber imperfecciones en el piso o imprecisión en sus motores, etc. Los MDPs consideran la incertidumbre en el resultado de las acciones del robot; si también existe incertidumbre sobre el estado actual del robot (que tiene que ver con su percepción del ambiente) existe otro modelo llamado MDP parcialmente observable (POMDP). Mediante MDPs y POMDPs se pueden modelar los estados y acciones de un robot para realizar cierta tarea de forma que su comportamiento sea más robusto.

Los MDPs se pueden utilizar para representar y resolver problemas robóticos a diferentes niveles, tanto muy básicos como para desplazarse en su ambiente, como de alto nivel para coordinar sus diferentes capacidades y resolver tareas complejas.¹⁵ En esta sección nos enfocamos a este último aspecto. Para realizar una tarea compleja un robot requiere utilizar diferentes capacidades en el nivel de dispositivos y algoritmos. Se requiere además que dichas capacidades se ejecuten de una forma coordinada para realizar una tarea, de preferencia en forma óptima (menor tiempo, mayor efectividad, etc.).

Mediante MDPs (o POMDPs) se pueden coordinar las diferentes capacidades básicas de un robot para realizar tareas complejas bajo incertidumbre. Para ello se define un *modelo* que especifica:

1. Un conjunto de estados \mathcal{S} que define la situación actual del robot a alto nivel; por ejemplo su ubicación, la meta, lo que conoce del ambiente, los comandos de sus usuarios, etcétera.
2. Un conjunto de acciones \mathcal{A} , que establece las posibles decisiones que puede realizar el robot, como localizarse, desplazarse de un lugar a otro, percibir el ambiente con sus cámaras u otros sensores, preguntar algo, etcétera.

¹⁵ Elinas, P., Sucar, L.E., Reyes, A., Hoey, J. (2004). **A decision theoretic approach for task coordination in social robots.** En *IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN)*, pp. 679–684.

3. Una función de transición $P(S_j, A, S_i)$ que define las probabilidades de que el robot pase a un nuevo estado S_j dado que realizó cierta acción A y se encontraba en cierto estado S_i . Estas probabilidades se pueden establecer en forma aproximada por una persona (subjetivas) o se pueden aprender a través de interactuar con el ambiente mediante algoritmos de aprendizaje por refuerzo.
4. Una función de recompensa que establece que tan deseable es cada estado (o estado-acción) de acuerdo a los objetivos de la tarea. Normalmente se establece una recompensa alta para el estado meta de la tarea; por ejemplo, el que el robot entregue a la persona cierto objeto, si la tarea es que el robot vaya y busque dicho objeto, lo tome y se lo lleve a una persona.

Una vez establecido el modelo de una tarea mediante un MDP existen algoritmos que permiten establecer una *política* o plan general para que el robot realice la tarea de forma óptima, es decir que maximice la utilidad esperada, que normalmente es la suma de las recompensas que recibe de acuerdo a la función establecida previamente. La política establece para cada estado cuál es la mejor acción que debe hacer el robot de forma que se lleve a cabo la tarea de la mejor manera, incluso habiendo incertidumbre en los resultados de sus acciones.

Este esquema permite que se pueda hacer un desarrollo modular y eficiente de robots de servicio para realizar diferentes tareas. Si se cuenta con un conjunto de capacidades básicas, desarrollar nuevas tareas implica simplemente cambiar el modelo del MDP y resolverlo para tener el plan correspondiente. Una analogía sería el del director de una orquesta que con los mismos músicos e instrumentos puede interpretar diferentes obras cambiando la partitura.

Existen diversas extensiones al esquema básico descrito anteriormente. Una es considerar incertidumbre en el estado actual del robot de forma que

se modela como un POMDP. Esto hace más compleja la solución ya que hay que considerar la probabilidad de cada estado (estado de creencia), lo que en principio implica un número infinito de estados; la solución exacta es sólo posible para problemas muy pequeños pero se han desarrollado soluciones aproximadas que permiten resolver casos más complejos como los de robótica. Otra extensión es el considerar que el robot puede realizar varias acciones a la vez, como desplazarse, comunicarse y reconocer personas. Para esto se han desarrollado MDPs concurrentes que toman en cuenta las restricciones para evitar conflictos entre los diversos tipos de acciones.¹⁶

El robot Markovito,¹⁷ el cual se ilustra en la Figura 4.2 (c), utiliza MDPs para representar y realizar diferentes tareas, entre las cuales están: i) llevar mensajes y objetos entre personas; ii) buscar un objeto en un ambiente doméstico; iii) desplazarse a diferentes cuartos en una casa; iv) recibir y reconocer a diferentes personas, entre otras.

4.6. Nivel de dispositivos y algoritmos

Este nivel se organiza en relación a las modalidades principales de la percepción, como la visión y el lenguaje así como de la acción motora incluyendo la navegación y manipulación. Cada una de estas facultades se habilita mediante sensores y actuadores físicos en conjunto con algoritmos especializados que interpretan la señal proveniente del entorno para construir una representación del mundo. Es importante recalcar que el robot no mantiene representaciones del mundo directamente sino de las interpretaciones del mundo que es capaz de crear con su equipamiento sensorial y sus procesos perceptuales. A continuación se describen las facultades robóticas principales que se han estudiado por la comunidad mexicana en i) percepción y acción motora, ii) audición y lenguaje y iii) representación del conocimiento y razonamiento.

¹⁶ Corona, E., Morales, E., Sucar, L.E. (2009). **Executing Concurrent Actions with Multiple Markov Decision Processes**. En *IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning*. IEEE Press, pp. 82–89.

¹⁷ <http://ccc.inaoep.mx/~markovito/>

4.6.1. Percepción y acción motora

La percepción robótica involucra dos tareas principales: i) el sentido de una señal proveniente del mundo mediante algún tipo de dispositivo, y ii) asignar una interpretación a dicha señal de manera coherente con el contexto o situación en que se hace la observación. Las señales pueden ser de cualquier modalidad de la percepción. Normalmente se presentan al sistema de cómputo como arreglos o matrices de datos, donde cada celda contiene un número que codifica a una propiedad de la señal sensada; por ejemplo, en el caso de las cámaras, la escena fotografiada se registra en una matriz de píxeles, donde cada celda contiene un número que representa el color del punto correspondiente en el espacio; a partir de estas informaciones se inicia el proceso de interpretación.¹⁸ Entre los tipos de sensores más comúnmente utilizados en robótica se encuentran los siguientes:¹⁹

- Sonares: emiten una señal sonora y a partir del tiempo de ida y vuelta se puede estimar la distancia a los obstáculos (análogo a cómo los murciélagos detectan obstáculos).
- Láser: emiten rayos láser y también a partir del tiempo de ida y vuelta (o la fase) se estima la distancia a los objetos; son más precisos que los sonares.
- Cámaras:²⁰ obtienen imágenes/videos del entorno que pueden utilizarse para múltiples propósitos.
- Cámaras de profundidad: como el Kinect, no sólo obtienen la información de color de cada píxel de la imagen sino también un estimado de su distancia.

¹⁸ Ver capítulo 7.

¹⁹ Estos tipos de dispositivos de alta tecnología se desarrollan en México sólo de manera muy limitada; por lo mismo, la investigación en robótica en nuestro país se hace comúnmente con componentes básicos disponibles en el mercado internacional.

²⁰ Dado su bajo costo, tamaño y versatilidad, las cámaras y las cámaras de profundidad se han vuelto muy populares recientemente. En muchos casos se combinan diferentes sensores, ya sea del mismo o diferente tipo, mediante lo que se conoce como *fusión sensorial*.

- Micrófonos: permiten percibir los sonidos en el ambiente, en particular la voz humana para la interacción con las personas.
- Narices electrónicas: recientemente se han desarrollado sensores olfativos que permiten detectar ciertos compuestos como el alcohol.²¹

Por su parte, la acción motora consiste en moverse en el entorno, caminar o tomar objetos con las manos, entre muchas otras formas. Estos procesos son inversos con respecto a la percepción ya que su entrada es una señal producida por un proceso computacional y la salida es la acción mecánica producida por un servo-motor.²²

No todas las tareas robóticas requieren que la escena se interprete completamente y es común que la información proporcionada por el sensor sólo permita una interpretación muy limitada; sin embargo, ésta puede ser suficiente para producir la conducta motora. Por ejemplo, para evadir objetos durante la navegación es suficiente detectar que hay un objeto enfrente para que el robot cambie de trayectoria sin tener que saber cuál es el objeto particular que se evade o a qué clase pertenece. Conductas de este tipo son reactivas en oposición a aquellas que involucran un proceso de interpretación profundo, que son deliberativas. Dentro de la robótica clásica, la mayoría de las conductas son reactivas e involucran al menos un sensor y un dispositivo de acción, por lo que la percepción y la acción motora se tienen que abordar juntas. Algunas de las tareas paradigmáticas de carácter reactivo son la navegación, el seguimiento o *tracking* y el movimiento supervisado por la visión o control visual. Por su parte, si el proceso incluye interpretaciones e inferencias deliberativas es posible diferenciar más claramente a las tareas de la percepción de las acciones motoras y estudiar ambos problemas por separado. Algunos procesos con un carácter más perceptual son la detección y el reconocimiento de personas, el reconocimiento de objetos y el análisis de escenas.

²¹ <http://tecreview.itesm.mx/esta-nariz-robotica-podria-salvarte-la-vida/>

²² Los servo-motores de calidad se producen también por corporativos internacionales y en México se usan como componentes básicos.

También hay tareas que involucran la construcción de representaciones pero que a su vez están estrechamente vinculadas a conductas motoras, por lo que la percepción y la conducta motora se abordan también en conjunto. Un ejemplo es la construcción de mapas al tiempo que se explora el espacio, aunada a la localización del robot en el mismo mapa; otro es el reconocimiento visual y la manipulación de objetos. Para enfrentar estos retos se han desarrollado una diversidad de algoritmos, algunos más genéricos, por ejemplo en el campo de la visión computacional y otros más específicos a la robótica. A continuación se mencionan algunos desarrollos de la comunidad mexicana en i) mapeo, auto-localización y navegación, ii) seguimiento, iii) reconocimiento y manipulación de objetos, iv) detección y reconocimiento de personas y v) monitoreo y vigilancia.

4.6.1.1. Mapeo, auto-localización y navegación

Uno de los problemas fundamentales para los robots de servicio es la movilidad, lo que incluye la construcción de mapas, la localización y la navegación:

1. Construcción de mapas: los robots móviles requieren una representación de su entorno o mapa para moverse en el entorno espacial; en general es conveniente que ellos mismos lo construyan por medio de sus capacidades sensoriales y algorítmicas. Existen diferentes tipos de mapas, incluyendo los métricos (rejillas de ocupación espacial, representaciones geométricas, etc., en 2 y 3 dimensiones), los topológicos, los semánticos²³ y los de configuraciones (que tienen tantas dimensiones como los grados de libertad del robot).
2. Localización: este problema consiste en que el robot determine su ubicación en el mapa de manera dinámica. La localización tiene dos variantes principales: i) global, en la cual el robot se tiene que localizar de manera absoluta en relación al mapa (por ejemplo cuando no sabe

²³ Ver sección 4.6.3.

donde está al iniciar una tarea); ii) local o relativo a una posición de referencia durante su desplazamiento. Para localizarse, el robot puede utilizar información distintiva del ambiente, como marcas (naturales o artificiales), objetos específicos o puntos característicos (esquinas, SIFT,²⁴ etc.). Es importante destacar que el robot tiene que actualizar su posición continuamente ya que si sólo utiliza el registro de su propio movimiento, lo que se conoce como odometría, el error de localización se incrementa continuamente hasta que eventualmente el robot se pierde.

3. SLAM: frecuentemente la construcción del mapa y el cálculo de la localización o ubicación se tienen que realizar de manera simultánea; este proceso coordinado se conoce como *mapeo y localización simultáneos* (SLAM, por sus siglas en inglés).
4. Navegación: Los robots requieren construir un plan o definir una trayectoria para desplazarse de un lugar a otro; esto se conoce como planeación de trayectorias o de movimiento (*motion planning*). Para esta tarea se han desarrollado diversos algoritmos, incluyendo determinísticos y probabilísticos. Una vez establecido, el plan debe ejecutarse, lo que involucra el control de los actuadores del robot para poder seguir la trayectoria deseada, como las llantas en robots móviles, las articulaciones en brazos robóticos y las piernas en robots humanoides. En ambientes variables deben detectarse posibles cambios en el mundo u obstáculos dinámicos y evitarlos o ajustar el plan.

Para resolver estas tareas en el INAOE y el CIMAT se han desarrollado algoritmos probabilísticos que utilizan sensores láser para obtener una representación del espacio libre (en 2-D) basados en rejillas de ocupación espacial, donde cada celda contiene un número real que representa la probabilidad

²⁴ SIFT (*Shift Invariant Feature Transform*) es uno de múltiples esquemas para identificar puntos característicos salientes y estables en imágenes: https://en.wikipedia.org/wiki/Scale-invariant_feature_transform.

de que esté libre u ocupada; con base al mapa resultante se han desarrollado técnicas de navegación robustas ante las limitaciones perceptuales del robot.²⁵ Asimismo, dicho mapa se utiliza para que el robot se localice mediante la detección de marcas naturales en el ambiente, por ejemplo esquinas, puertas y paredes.²⁶ También se ha trabajado en técnicas de navegación mediante el aprendizaje de reglas tele-reactivas a partir de ejemplos provistos por una persona.²⁷

Se ha trabajado además en la navegación de robots con ruedas basada en un conjunto de imágenes objetivo²⁸ las cuales se capturan previamente y constituyen una representación del ambiente que se denomina “memoria visual”, la cual le permite al robot navegar posteriormente de forma autónoma. Se tienen ya resultados para robots humanoides. Otro reto es hacer mapas utilizando una sola cámara, lo que se conoce como SLAM monocular, ya que no se tiene información de profundidad; sin embargo, ésta se puede estimar mediante conocimiento previo del mundo.²⁹ Un reto más, abordado en el CINVESTAV, IPN y el ITAM es la navegación para coches autónomos que requiere la integración a alta velocidad de componentes de localización basada en entrada visual y de lidiar con planificación de tareas y de movimientos.

²⁵ Romero, L., Morales, E., Sucar, L.E. (2001). **A robust exploration and navigation approach for indoor mobile robots merging local and global strategies.** En *International Conference on Robotics and Automation (ICRA)*, IEEE Press, pp. 3092–3097.

²⁶ Romero, L., Morales, E., Sucar, L.E. (2001). **A Hybrid Approach to Solve the Global Localization Problem for Indoor Mobile Robots Considering Sensor’s Perceptual Limitation’s.** En *Int. Joint Conf. on Artificial Intelligence (IJCAI)*, AAAI Press, pp. 1411–1416.

²⁷ Vargas, B., Morales, E. (2009). **Learning Navigation Teleo-Reactive Programs using Behavioural Cloning.** En *Proc. IEEE International Conference on Mechatronics (ICM)*.

²⁸ Becerra, H. M. , Sagues, C., Mezouar, Y., Hayet, J. B. (2014). **Visual navigation of wheeled mobile robots using direct feedback of a geometric constraint.** *Autonomous Robots*, 37(2): 137–156.

²⁹ Mota-Gutierrez, S. A., Hayet, J. B., Ruiz-Correa, S., Hasimoto-Beltran, R., Zubieta-Rico, C. E. (2015). **Learning depth from appearance for fast one-shot 3-D map initialization in VSLAM systems.** En *Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA)*, pp. 2291–2296.

Un problema relacionado es la caminata estable de robots humanoides. Se han hecho desarrollos para simultáneamente realizar el control visual y controlar la caminata dinámica de un robot humanoide.³⁰ Para este efecto se utiliza un generador de patrón de movimiento basado en control predictivo (MPC) que adapta los pasos de la caminata y las trayectorias del centro de masa del robot para seguir perfiles de velocidad provistos por el usuario. La contribución de este trabajo es de reformular el problema de MPC al considerar información de retroalimentación visual, en lugar de usar la velocidad de referencia. Un problema relacionado es la regulación de postura por medio de visión para robots humanoides.³¹ Para este efecto se utiliza un esquema de control capaz de lograr la convergencia de la tarea visual en un tiempo predefinido. El esquema es un control jerárquico basado en tareas que puede manejar al mismo tiempo la tarea visual y una tarea de evasión de obstáculos, los cuales se detectan con una cámara de profundidad. Genera velocidades continuas para el robot y considera las capacidades de caminata omnidireccional de los robots humanoides.

4.6.1.2. Seguimiento

El seguimiento de personas en ambientes dinámicos donde pueden transitar libremente vehículos y personas, como en las pruebas de navegación de la competencia RoboCup, es una tarea sumamente compleja. Este problema se ha abordado por los robots Markovito, Justina y Golem con diferentes estrategias y algoritmos.

El método de seguimiento de personas del equipo Markovito del INAOE combina información del esqueleto obtenido con el Kinect con información

³⁰ Garcia, M., Stasse, O., Hayet, J. B., Dune, C., Esteves, C., Laumond, J. P. (2015). **Vision-guided motion primitives for humanoid reactive walking: Decoupled versus coupled approaches.** *International Journal of Robotics Research*, 34(4-5): 402–419.

³¹ Delfin, J., Becerra, H. M., Arechavaleta, G. (2016). **Visual servo walking control for humanoids with finite-time convergence and smooth robot velocities.** *International Journal of Control*, 89(7): 1342–1358.

de color de la cámara del robot. Inicialmente se detecta a la persona mediante el algoritmo de detección de esqueletos del Kinect y luego se aprende un modelo de color de la ropa de dicha persona. Para el seguimiento se integra la información de las observaciones (basadas en color) con la estimación de movimiento mediante un filtro de Kalman.

La estrategia que se desarrolla actualmente por el Grupo Golem se basa en el uso combinado de un láser que localiza el torso de la persona a seguir y el esqueleto que se obtiene del Kinect. El método utiliza adicionalmente el parche de colores del objeto a seguir, como la camisa o el pantalón del sujeto, para distinguirlo de otras personas u objetos, fijos o móviles, que se encuentren en la trayectoria de seguimiento.

Un problema relacionado que se considera clásico en robótica móvil desde el punto de vista de los sistemas de control es el de persecución–evasión, donde un robot debe seguir y alcanzar a un blanco móvil. Este problema se ha analizado en el CIMAT desde el punto de vista teórico para la captura de un evasor omnidireccional usando un robot de manejo diferencial en un ambiente sin obstáculos.³² El método propuesto utiliza técnicas de control óptimo para obtener representaciones de las primitivas de movimiento y estrategias (en equilibrio de Nash) de tiempo mínimo para seguidor y evasor. También aborda el problema de decisión del juego y las condiciones que definen al ganador. Esta metodología se ha aplicado además al problema de persecución–evasión en un ambiente con obstáculos,³³ con especial atención al problema combinatorio que surge para mantener visibilidad de un evasor que visita varias ubicaciones.

³² Ruiz, U., Murrieta-Cid, R., Marroquín, J. L. (2013). **Time-optimal motion strategies for capturing an omnidirectional evader using a differential drive robot.** *IEEE Transactions on Robotics*, 29(5): 1180–1196.

³³ Becerra, I., Murrieta-Cid, R., Monroy, R., Hutchinson, S., Laumond, J. P. (2016). **Maintaining strong mutual visibility of an evader moving over the reduced visibility graph.** *Autonomous Robots*, 40(2): 395–423.

4.6.1.3. Reconocimiento y manipulación de objetos

Los robots de servicio deben ser capaces de tomar y llevar objetos de un sitio a otro, lo que se conoce como el problema de manipulación. Para ello los robots móviles y humanoides incorporan brazos robóticos con pinzas o manos para sujetar diferentes tipos de objetos. La manipulación presupone a su vez que el robot debe ver los objetos para lo cual es necesario identificarlos y determinar su posición y orientación espacial. Por esta razón los problemas de reconocimiento y manipulación se relacionan de manera muy estrecha. Desde el punto de vista computacional el problema de manipulación involucra varios subproblemas:

- Reconocimiento y ubicación del objeto: el primer paso para manipular un objeto es reconocerlo y ubicar su posición en el ambiente (coordenadas 3-D).
- Modelado: en el caso de objetos complejos se requiere conocer su forma; es decir, obtener un modelo tridimensional que permita identificar posibles puntos de agarre. Para ello se obtienen diferentes vistas del objeto con algún sensor de profundidad, como el Kinect, que se integran en una representación 3-D. El determinar las mejores posiciones del sensor para obtener dichas vistas es un problema de planificación que se conoce como “siguiente mejor vista” (*next best view*).
- Determinar puntos de sujeción (*grasping*): en base al modelo 3-D del objeto se deben establecer puntos de sujeción estables (normalmente dos o tres, dependiendo del tipo de pinza o mano); es decir, que permitan sujetar y levantar al objeto en forma segura.
- Planeación/ejecución de movimientos: una vez que se establece la posición y punto de agarre del objeto es necesario llevar al efector final del brazo a las coordenadas deseadas, para lo cual se establece un plan de los movimientos del brazo/base móvil que luego se ejecuta,

normalmente con cierta retroalimentación sensorial que asegure tomar al objeto.

Enseguida resumimos algunas aportaciones de grupos en nuestro país.

El Grupo Golem del IIMAS, UNAM, aborda la tarea de reconocimiento de objetos mediante el algoritmo MOPED³⁴ el cual requiere crear un modelo de antemano para cada objeto conocido. Los modelos se forman a partir de un conjunto de fotografías tomadas alrededor del objeto desde diferentes distancias y perspectivas para configurar una envolvente. Las fotografías se procesan mediante el algoritmo SIFT.³⁵ Con estos descriptores se forma una nube de puntos alrededor del objeto, la cual se calibra con respecto a un punto externo de observación. Esta calibración permite determinar cuál será el tamaño del objeto con relación a un nuevo punto de observación. Para efectos de reconocimiento, la vista del objeto desde el punto de observación se procesa también mediante el algoritmo SIFT y sus descriptores se alinean parcialmente con respecto a alguna vista arbitraria del modelo. Si este proceso es exitoso el objeto se reconoce en conjunto con su posición y orientación o *pose* en el espacio 3-D en relación al punto de observación, que corresponde con la cámara del robot.

Para efectos de la manipulación el robot alinea su brazo a la misma altura que el objeto y, dado que conoce el tamaño de su brazo y la distancia al objeto, le es posible dirigir su brazo hacia la posición del objeto por un método de triangulación directo. Con esta metodología el equipo Golem obtuvo el 2º lugar con el robot Golem-II+ en el reto técnico de la competencia RoboCup México 2012 y el 3er. lugar en la prueba de visión y manipulación en el RoboCup Leipzig 2016 con el robot Golem-III. En esta misma competencia Golem-III se desempeñó como un asistente de supermercado en la prueba

³⁴ Collet, A., Martínez, M., Srinivasa, S.S. (2011). **The MOPED framework: Object Recognition and Pose Estimation for Manipulation.** *The International Journal of Robotics Research*, 30: 1284–1306.

³⁵ Ver nota 23.

del *Open Challenge* en la que tomó varios objetos exitosamente y obtuvo el 2° lugar.

El algoritmo MOPED es muy preciso (tiene muy pocos falsos positivos) pero sólo funciona con objetos ricos en textura, además de que no puede reconocer objetos para los que no se cuente con modelos, por lo que su cobertura es pobre y puede tener muchos falsos negativos (es decir objetos que no ve). Para superar esta limitación el grupo Golem utiliza una segunda estrategia basada en la nube de puntos, que corresponden a objetos en la escena, que se obtiene a través del sensor de distancia del Kinect. Las nubes se segmentan en la imagen y se asocian a una máscara de color, la cual se utiliza como un identificador único para objetos de baja textura. Con esta nueva funcionalidad, el robot Golem-III pondera los objetos que reconoce y otorga un mayor peso a los objetos reconocidos con MOPED, un peso menor a los objetos modelados por nube de puntos y máscara de color, y un peso aún más bajo a los objetos que reconoce por la nube de puntos, pero cuya máscara de color no corresponde con ninguno de los modelos conocidos. Estos últimos objetos se catalogan como “no conocidos”. Con esta estrategia en cascada se espera incrementar significativamente la cobertura visual y dar al robot Golem-III la capacidad de manipular una gama amplia de objetos tanto conocidos como desconocidos.

Sin embargo, es difícil que un robot de servicio cuente con modelos previos de todos los posibles objetos en el mundo, por lo que en el INAOE se ha desarrollado un método de reconocimiento de objetos que aprende nuevos modelos a partir de ejemplos de imágenes que obtiene de Internet.³⁶ Para tener más ejemplos se generan nuevas imágenes mediante transformaciones de las imágenes originales y luego se aprende un clasificador del nuevo objeto a partir de los ejemplos, que se puede usar para reconocerlos en el ambiente.

³⁶ Navarrete, D., Morales, E., Sucar, L. E. (2012). **Unsupervised learning of visual object recognition models**. En *IBERAMLAI - LNAI 7637*, Springer-Verlag, pp. 511–520.

En el INAOE se desarrolló también un algoritmo para resolver el problema de planificación de vistas para modelado tridimensional de objetos; éste toma en cuenta la incertidumbre en la posición final del sensor, el cual está montado en el eslabón final de un brazo sobre un robot móvil.³⁷ El desplazamiento de la base móvil introduce errores en la posición deseada del sensor, lo que puede ocasionar que no se obtenga una vista adecuada. Para compensar esto y tener posiciones robustas ante el error de posicionamiento, el planificador ejecuta varias veces el algoritmo de planeación basado en RRT³⁸ (*Rapid exploring Random Tree*) simulando errores y selecciona aquel plan que da un valor esperado óptimo. Para ello se define una función de utilidad que establece un compromiso entre varios aspectos deseables: i) observar superficies no vistas del objeto, ii) garantizar un traslape mínimo con las vistas anteriores (para poder juntar las vistas en un modelo 3-D), y iii) minimizar la distancia que el robot debe moverse.

En el campo del control visual de robots móviles se han propuesto esquemas genéricos de control robusto que usan restricciones geométricas y relacionan múltiples vistas de una escena. Por ejemplo, el problema de regulación de postura se ha abordado mediante un esquema de control visual basado en imágenes (*image-based visual servoing*) que usa restricciones geométricas para conducir a un robot de manejo diferencial a una posición y orientación deseadas.³⁹ Un problema relevante en este contexto es el seguimiento de objetos, por ejemplo, para tomar un objeto con la mano. En el CIMAT se ha investigado también el problema de confirmar la identidad de un objeto candidato; para este efecto se propuso un método que mezcla la localización del robot relativa al objeto candidato y la confirmación de que dicho objeto es el

³⁷ Vasquez-Gomez, J.I., Sucar, L.E., Murrieta-Cid, R. (2016). **View/state planning for three-dimensional object reconstruction under uncertainty.** *Autonomous Robots*, 41 (1):89–109.

³⁸ https://en.wikipedia.org/wiki/Rapidly-exploring_random_tree

³⁹ Becerra, H. M., Hayet, J. B., Sagués, C. (2014). **A single visual-servo controller of mobile robots with super-twisting control.** *Robotics and Autonomous Systems*, 62 (11): 1623–1635.

blanco buscado.⁴⁰ El proceso de confirmación con esta meta dual se modela como un POMDP que se resuelve con Programación Dinámica.

Otro problema relevante es la búsqueda de objetos; por ejemplo si el usuario hace una petición (robot: “tráeme una manzana”) en una casa es necesario combinar aspectos de percepción, planeación y navegación. Se desea que el robot encuentre el objeto en el menor tiempo posible, para lo cual debe estimar los lugares más probables (por ejemplo, una manzana en la cocina o el comedor) y establecer una estrategia de búsqueda que minimice el tiempo esperado. En el INAOE se desarrolló una estrategia que inicialmente estima la probabilidad de encontrar cierto objeto en los diferentes tipos de habitaciones; ésta se basa en estimar la correlación entre objeto-habitación mediante el uso de diversas bases de datos en Internet.⁴¹ Combinando la probabilidad de las habitaciones, la distancia del robot a las habitaciones y el área de las habitaciones se establece una heurística que permite planear una ruta para explorar las habitaciones en el menor tiempo posible, la cual da resultados cercanos al óptimo.^{42, 43} Finalmente, el robot busca el objeto dentro de cada habitación enfocándose en superficies planas donde se espera encontrarlo.

4.6.1.4. Reconocimiento de personas

El reconocimiento de personas forma parte de las tareas que un robot de servicio debe realizar. Este problema se puede descomponer en tres subpro-

⁴⁰ Becerra, I., Valentín-Coronado, L. M., Murrieta-Cid, R., Latombe, J. C. (2016). **Reliable confirmation of an object identity by a mobile robot: A mixed appearance/localization-driven motion approach.** *International Journal of Robotics Research*, 35(10): 1207–1233.

⁴¹ Izquierdo-Córdova, R., Morales, E.F., Sucar, L.E. (2016). **Object Location Estimation in Domestic Environments through Internet Queries.** *RoboCup Symposium*, 2016

⁴² Izquierdo-Córdova, R., Morales, E.F., Sucar, L.E., Murrieta-Cid, R. (2016). **Searching Objects in Known Environments: Empowering Simple Heuristic Strategies.** *IJCAI Workshop on Service Robots*.

⁴³ Garantizar el tiempo óptimo es un problema difícil de resolver eficientemente, lo que se conoce en computación como problemas “NP” (no existe algoritmo conocido que pueda resolverlo en un tiempo polinomial respecto al tamaño del problema).

blemas: detección, localización e identificación, cuyo objetivo es determinar cuántas personas hay, dónde están y quiénes son respectivamente.⁴⁴

La detección está relacionada con la capacidad de clasificar cuáles objetos en el ambiente son personas y cuáles no. Una estrategia común es primero detectar caras en las imágenes;⁴⁵ para este efecto hay un algoritmo muy popular que utiliza técnicas de aprendizaje computacional para construir un ensemble de clasificadores⁴⁶ que permite determinar si un pixel forma parte de una cara o no. La localización, por su parte, consiste en determinar la ubicación exacta de la persona detectada. Por ejemplo, si la detección se realiza con Kinect es posible obtener las coordenadas 3-D de los puntos más relevantes de su cuerpo, incluyendo las manos, los codos, los hombros y la cabeza. Finalmente, la identificación consiste en determinar si la persona detectada está incluida en la base de datos del robot. Este último proceso se realiza principalmente mediante el uso de rasgos biométricos;⁴⁷ por ejemplo, mediante la identificación de puntos característicos en la zonas de los ojos-nariz-boca. El reconocimiento se hace más robusto al combinar varias imágenes mediante un esquema bayesiano,⁴⁸ que es la estrategia que usa el robot Markovito para reconocer personas.

En particular, el robot Golem-III realiza la detección de personas mediante el algoritmo de Viola-Jones para detectar rostros. También puede detectar la región completa de la cabeza y los hombros utilizando un algoritmo desarrollado por el grupo de investigación Golem, el cual se basa en la técni-

⁴⁴ Cielniak, G., Duckett, T. (2004). **People Recognition by Mobile Robots.** *Journal of Intelligent and Fuzzy Systems*, 15:21–27.

⁴⁵ Viola, P., Jones, M (2001). **Robust Real-time Face Detection.** *International Journal of Computer Vision*, 57(2):137–154.

⁴⁶ Ver capítulo 2.

⁴⁷ Jain, A.K., Ross, A., Prabhakar, S. (2004). **An Introduction to Biometric Recognition.** *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):4–20.

⁴⁸ Cruz, C., Sucar, L. E., Morales, E. (2008). **Real-Time Face Recognition for Human-Robot Interaction.** En *IEEE International Conference on Automatic Face and Gesture Recognition*.

ca de Histogramas de Gradientes Orientados.⁴⁹ El robot utiliza un Kinect 2 para poder detectar cuerpos y sus respectivos esqueletos, los cuales permiten clasificar la posición en la que se encuentra una persona –de pie, sentada o recostada en el piso– o determinar si está realizando un gesto –ya sea apuntar o alzar la mano para llamar la atención. Para la identificación de personas, el robot utiliza la técnica llamada Eigenrostros,⁵⁰ la cual es un enfoque biométrico que utiliza algoritmos de aprendizaje.

En años recientes, un tipo especial de rasgos llamados biométricos suaves⁵¹ (como el color de la ropa) han atraído la atención de distintas investigaciones debido a su utilidad en escenarios no controlados. En particular, el robot Golem-III puede identificar a una persona a partir de su ropa con la información que provee el detector de esqueletos de Kinect 2 y sin necesidad de ver el rostro. Para ello, durante la fase de aprendizaje, el robot guarda múltiples vistas de la misma persona en distintas orientaciones y extrae parches de color que se etiquetan semánticamente de acuerdo a la parte del cuerpo en la que se encuentran.

4.6.1.5. Monitoreo y vigilancia

Los robots no sólo deben ser capaces de analizar imágenes estáticas sino también secuencias de imágenes o videos. En el CIMAT se desarrolló una metodología para aprender y utilizar modelos de movimiento probabilísticos de múltiples objetivos, incorporada a un esquema de rastreo visual por filtro de partículas en secuencias de video.⁵² Dada una secuencias de video para

⁴⁹ Dalal, N., Triggs, B. (2005). **Histograms of oriented gradients for human detection.** *Computer Vision and Pattern Recognition*, 1:886-893.

⁵⁰ Turk, M., Pentland, A. (1991). **Eigenfaces for recognition.** *Journal of Cognitive Neuroscience*, 3(1):71–86.

⁵¹ Jain, A. K., Dass, S. C., Nandakumar K. (2004). **Soft Biometric Traits for Personal Recognition Systems.** En *Proc. of Inter. Conference on Biometric Authentication*, pp. 731-738.

⁵² Madrigal, F., Hayet J. B. (2015). **Goal-oriented visual tracking of pedestrians with motion priors in semi-crowded scenes.** En *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA)*, pp. 720-725.

entrenamiento se extraen primero las posiciones (u “objetivos”) donde los peatones entran/salen de la escena observada, o simplemente donde cambian de dirección con frecuencia. Posteriormente, se aprende un modelo de movimiento a partir de estadísticas para cada uno de estos objetivos. Finalmente, se utilizan los modelos de movimiento en un esquema de tipo IMM (Modelos Múltiples en Interacción) para el seguimiento y la estimación del objetivo de cada persona rastreada.

4.6.2. Audición y Lenguaje

Los robots de servicio requieren comunicarse a través del lenguaje, especialmente hablado. Para este efecto se deben incorporar sistemas de reconocimiento e interpretación del lenguaje así como de generación de lenguaje y síntesis de voz. Las contribuciones de la comunidad mexicana al reconocimiento de voz⁵³ se han utilizado directamente para habilitar el reconocimiento del español hablado a los robots producidos en el Proyecto Golem, especialmente a Golem-I y al sistema Golem en Universum. Asimismo, el lenguaje de programación SitLog se ha utilizado para crear los modelos de diálogo de los robots Golem-II+ y Golem-III. En estos robots, así como en Justina y Markovito, se ha utilizado también el sistema de reconocimiento de voz de Microsoft, especialmente para habilitar la comunicación en inglés en el contexto de las competencias RoboCup en la liga @Home,⁵⁴ en la que estos robots participan regularmente.

La producción del lenguaje, por su parte, tiene dos componentes principales: la generación de los contenidos y la síntesis de voz propiamente. Para el primero se han usado diferentes estrategias que van desde el uso de plantillas o “templates” hasta textos cuya generación está asociada a un proceso inferencial; para el segundo se utilizan sintetizadores con voces sintéticas disponibles en Internet, ya que no hay productos locales lo suficientemente robustos para estas tareas.

⁵³ Ver capítulo 3.

⁵⁴ <http://www.robocupathome.org/>

4.6.2.1. Audición robótica y análisis de imágenes acústicas

El reconocimiento de voz en robots de servicio presenta tres retos adicionales a lo descrito anteriormente: i) detección y selección de varias fuentes sonoras, por ejemplo varias personas que interactúan al mismo tiempo con el robot, ii) separación de estas fuentes sonoras, y iii) su clasificación.

El problema de la localización de fuentes sonoras se ha abordado con una configuración de hardware, con un arreglo de tres micrófonos dispuestos en un triángulo equilátero que presenta tres frentes a los sonidos generados en el entorno, de tal forma que la diferencia de fases entre cada par de micrófonos en cada frente permite establecer una hipótesis acerca de la posición de la fuente sonora. Asimismo, la posición de la fuente en relación a los tres frentes del arreglo debe ser lo más coherente posible, por lo que las hipótesis incoherentes se descartan directamente. Este arreglo, en conjunto con su algoritmo asociado, permite localizar más fuentes que el número de micrófonos utilizados, superando las limitantes de los sistemas tradicionales.⁵⁵ El único requisito es que las fuentes sonoras sean de voz humana ya que el algoritmo aprovecha que si dos o más personas hablan a la vez, el empalme de la voz en micro segmentos de tiempo (por ejemplo del orden de 100 milisegundos) es muy bajo. En los momentos de poco empalme se detecta a una persona y, a lo largo del tiempo, se agrupan las detecciones por cada fuente. Este sistema es muy liviano en requisitos de hardware y software, por lo que es apropiado para aplicaciones de robótica de servicio. Dicho desarrollo fue integrado en el robot de servicio Golem y se utiliza en varias aplicaciones como⁵⁶ recuperarse al perder visualmente a la persona que se sigue, tomar asistencia en una clase, jugar el juego Marco Polo (versión americana del juego “Gallinita Ciega”)

⁵⁵ Rascon, C., Fuentes, G., Meza, I. (2015). **Lightweight multi-DOA tracking of mobile speech sources.** *EURASIP Journal on Audio, Speech, and Music Processing*, 2015(11). doi:10.1186/s13636-015-0055-8.

⁵⁶ Meza, I., Rascón, C., Fuentes, G., Pineda, L. A. (2016). **On Indexicality, Direction of Arrival of Sound Sources, and Human-Robot Interaction.** *Journal of Robotics*, Article ID 3081048, 2016. doi:10.1155/2016/3081048.

y ser mesero en un restaurante. La aplicación de mesero⁵⁷ fue evaluada y se encontró que los usuarios sintieron que el robot Golem se comportaba como un mesero al ser llamado desde lejos. También se encontró una solución más natural al tomar la orden de varios comensales que hablan al mismo tiempo: cuando el robot detecta a más de una persona hablando, descarta la orden, pide que sólo una persona hable a la vez y, utilizando la información de dirección, encara a la persona que le toca hablar. Básicamente, el robot es el que impone el orden de la interacción.⁵⁸

Una vez ubicadas las fuentes sonoras es necesario separar las voces de personas que estén hablando al mismo tiempo. Para este efecto es necesario separar las señales y alimentar la información de un sólo hablante a la vez al reconocedor de voz. Adicionalmente, si una de las fuentes es ruido (por ejemplo, tráfico en la calle, un abanico de aire, música, etc.) se filtra de manera natural. Más aún, el robot puede ser capaz de reconocer a dos personas que estén hablando al mismo tiempo. Actualmente, dicho reto se aborda en la UNAM por medio de técnicas de *Aprendizaje Profundo*.

El último paso de este proceso es clasificar las fuentes sonoras; si una fuente es una persona y la otra ruido se puede omitir alimentar la información del ruido al reconocedor y hacer más eficiente el reconocimiento de voz. Adicionalmente, si varias personas están hablando a la vez se puede identificar quién está hablando y, posiblemente, darle más prioridad al dueño del robot que a un tercero.

⁵⁷ Rascon, C., Meza, I., Fuentes, G., Salinas, L., Pineda, L. A. **Integration of the Multi-DOA Estimation Functionality to Human-Robot Interaction.** *Int J Adv Robot Syst*, 2015, 12:8. doi: 10.5772/59993

⁵⁸ Esta aplicación se demostró como la prueba *Open Challenge* del equipo Golem en el *RoboCup Eindhoven 2013*, en Holanda, en la cual se obtuvo la puntuación más alta y se premió con el *Innovation Award* de dicha competencia.

El corpus DIMEx100⁵⁹ (descrito en el capítulo 3) se ha utilizado también para la evaluación de desarrollos de Audición Robótica en México y es la base del corpus de evaluación *Acoustic Interactions for Robot Audition* (AIRA)⁶⁰ que simula fuentes sonoras por medio de las mismas grabaciones de DIMEx100 reproducidas en localizaciones conocidas, con múltiples micrófonos con posiciones conocidas, otorgando transcripciones y señales originales limpias.

4.6.3. Representación del conocimiento y razonamiento

En robótica de servicio es muy importante contar con un sistema de representación de conocimiento y razonamiento. Entre los diversos aspectos que debe conocer el robot destaca la necesidad de contar con un mapa del entorno físico. Los principales tipos de mapas utilizados en robótica son:

1. *Mapas métricos.* Consisten en una representación geométrica del espacio en dos o tres dimensiones en la que se distinguen al menos las zonas libres (por donde el robot puede desplazarse) de las zonas ocupadas (obstáculos). Esto le permite al robot conocer su ubicación en el mundo y navegar. Existen diversas formas de representación incluyendo: i) mapas de celdas (voxels en 3-D) en el cual el espacio se divide en una rejilla de 2 ó 3 dimensiones a cuyas celdas se les asigna un valor de ocupado o libre (cuantificado por una probabilidad); ii) mapas que se representan con elementos geométricos básicos, como líneas, superficies, cilindros, etc., que al integrarse proveen un modelo aproximado del mundo; iii) mapas en base a puntos característicos que se extraen de las imágenes y cuyo objetivo principal es permitir al robot mantener su ubicación, sin incluir un modelado preciso de los objetos en el mundo.

⁵⁹ Pineda, L. A., Castellanos, H., Cuétara J., Galescu, L., Juárez, J., Llisterri, J., Pérez P., Villaseñor, L. (2010). **The Corpus DIMEx100: Transcription and Evaluation.** *Language Resources and Evaluation* 44:347–370. doi: 10.1007/s10579-009-9109-9

⁶⁰ Ver nota 15.

2. *Mapas topológicos*. El ambiente se representa en forma más abstracta en base a un grafo de conectividad, donde los nodos representan áreas o espacios (como cuartos en ambientes interiores) y los arcos denotan relaciones de adyacencia o conectividad (como pasillos y puertas). Los mapas topológicos le permiten al robot planear sus trayectorias sin necesidad de información precisa del mapa.
3. *Mapas de configuraciones*. Representan los movimientos posibles que puede realizar un robot. Una configuración es una abstracción del robot que representa en forma única su pose y tiene un parámetro por cada uno de sus grados de libertad. Por ejemplo, un robot móvil sin manipulador tiene usualmente tres, dos para desplazamiento y uno para orientación, y una configuración es una combinación única de estos tres parámetros. Estos mapas se representan como grafos cuyos nodos son configuraciones y sus aristas secuencias de configuraciones para transitar entre los nodos. La construcción de estos mapas es intratable computacionalmente por lo que se obtienen mediante muestreo ya sea en forma global (como los *Probabilistic Roadmaps* o PRMs) o incremental (como los *Rapidly Exploring Random Trees* o RRTs)
4. *Mapas semánticos*. Expresan información de más alto nivel, como el tipo de ambiente (interiores, exteriores) o la clase de habitación (sala, comedor, recámara, etc.) que permiten dar comandos en forma más natural a los robots, por ejemplo: “robot, tráeme una manzana de la cocina”.

La construcción del mapa por el propio robot permite que la representación sea coherente con sus capacidades sensoriales; esto ha motivado un desarrollo significativo de las técnicas de SLAM, las cuales se estudian en México desde los noventa.⁶¹

⁶¹ Por ejemplo, Romero, L., Sucar, L.E., Morales, E. (2000). **Learning Probabilistic Grid-Based Maps for Indoor Mobile Robots Using Ultrasonic and Laser Range Sensors**. En *MICAI 2000 - LNAI 1793*, Springer-Verlag, pp. 158–169.

Además de contar con conocimiento del mundo físico en el que habita, el robot puede tener otros tipos de conocimiento, como:

- Conocimiento sobre los objetos en su ambiente y sus propiedades, incluyendo objetos pequeños (manipulables), objetos fijos, etcétera.
- Conocimiento sobre diversos agentes, como personas u otros robots (por ejemplo, para reconocerlas).
- Conocimiento de lenguaje que le permita comunicarse con personas y posiblemente con otros robots.

Para representar este conocimiento se utilizan los diversos formalismos que se han desarrollado en inteligencia artificial, incluyendo lógica de predicados, reglas de producción, representaciones estructuradas, modelos gráficos probabilistas, entre otros.

En particular el robot Golem-III cuenta con una base de conocimiento no-monotónica en la que se pueden representar clases de individuos en una estructura jerárquica, con sus propiedades y relaciones, así como individuos concretos de cada clase, también con sus propiedades y relaciones particulares.⁶² El sistema permite la expresión de “valores por omisión” o *defaults*, así como excepciones, por lo que permite algunas formas de razonamiento no-monotónico.⁶³ La base de datos de conocimiento se embebe en la conducta perceptual y motora del robot, y permite una gran expresividad y flexibilidad. El robot Golem-III cuenta también con un sistema de inferencia que le permite hacer diagnósticos, tomar decisiones (en relación a un conjunto de valores o preferencias preestablecidas) así como planes para satisfacer dichos objetivos. Por ejemplo, si en la tarea de asistente del supermercado Golem-III

⁶² Pineda, L. A., Rodríguez, A., Fuentes, G., Rascón, C., Meza, I. (2017). **A Light non-monotonic knowledge-base for service robots**. *Intel Serv Robotics* 10(3):159-171. doi:10.1007/s11370-017-0216-y. Ver también el capítulo 1.

⁶³ Idem.

no encuentra un objeto en el estante en el que debería estar es capaz de diagnosticar que el proveedor lo puso en un estante incorrecto, de tomar la decisión de ordenar los estantes y/o atender los requerimientos del cliente, y de construir un plan y llevarlo a cabo para lograr este objetivo.⁶⁴ Este ciclo de razonamiento se puede utilizar de manera directa como parte de la estructura de una tarea, o dinámicamente, cuando las expectativas del robot no se cumplen, por lo que se sale de contexto; en este caso el ciclo inferencial tiene por objetivo identificar qué pasó y cómo contextualizarse nuevamente para completar la tarea.

4.7. Nivel implementacional

Este nivel implementacional está constituido por la plataforma de hardware del robot, incluyendo las partes mecánicas y electrónicas, así como por los programas de apoyo que permiten el acceso a los dispositivos de hardware (sensores y actuadores) y la interacción entre programas y datos (sistema operativo). Presentamos primero algunas plataformas robóticas orientadas a robots de servicio, y luego los principales programas de apoyo, en particular los sistemas operativos robóticos.

4.7.1. Plataformas de robots de servicio

Las plataformas más comunes para robots de servicio son los robots móviles.⁶⁵ Éstas se desplazan sobre ruedas utilizando diversas estructuras (comúnmente la llamada “diferencial”) e incluyen en general los siguiente elementos:

- Motores para el control de las ruedas que permiten el desplazamiento del robot sobre el piso.
- Baterías que proveen la energía eléctrica a los diversos dispositivos.

⁶⁴ Ver nota 14.

⁶⁵ En la figura 2 se ilustran diversas plataformas robóticas de algunos grupos en México.

- Sensores internos que permiten conocer el estado del robot, como el nivel de la batería y su desplazamiento (odometría).
- Sensores externos para percibir su ambiente, como sonares, láser, infrarrojos, cámaras de video, cámaras de profundidad y micrófonos.
- Equipo de cómputo interno para el control de sensores, actuadores y la ejecución de los programas que proveen la funcionalidad al robot.
- Equipo de comunicación que permite conectarse a la red (Internet) y usar otro equipo de cómputo externo si es necesario.
- Brazos manipuladores (normalmente uno o dos) para poder tomar y transportar objetos.
- Otros elementos como el simular una cara que dé cierta apariencia humanoide al robot y le facilite la interacción con las personas.

La mayoría de estos dispositivos se importan en México, aunque ya hay compañías mexicanas que han logrado desarrollos de calidad de brazos, manos y torsos⁶⁶ así como desarrollos experimentales de cabezas y rostros con la capacidad de hacer gestos y expresar emociones, como alegría, tristeza o sorpresa.⁶⁷

⁶⁶ Por ejemplo, el Laboratorio de Innovación y Desarrollo Tecnológico (LAIDETEC), surgida del Departamento de Probabilidad y Estadística del IIMAS, UNAM, por iniciativa de Hernando Ortega Carrillo, con el apoyo de la Incubadora de Empresas de Base Tecnológica, InnovaUNAM. Sus líneas principales son el diseño y construcción de prótesis de mano robóticas y el desarrollo de una plataforma para robots de servicio con fines educativos y de investigación. Asimismo, los brazos, el torso, las pinzas y manos de Golem-III se diseñaron y construyeron por Hernando Ortega dentro del contexto de su participación y colaboración con el Grupo Golem.

⁶⁷ Por ejemplo la cabeza y rostro del robot Golem-III con la capacidad de expresar emociones se desarrollaron por Mauricio Reyes Castillo del Centro de Investigación en Diseño Industrial (CIDI) de la Facultad de Arquitectura de la UNAM en el contexto de su investigación doctoral en el Posgrado de Ciencia e Ingeniería de la Computación de la UNAM y su participación en el Grupo Golem.

Otra clase de robots de servicio son los robots humanoides que utilizan piernas robóticas para su desplazamiento. Tienen la ventaja de una mayor flexibilidad con respecto a los robots móviles ya que no están limitados a superficies planas. Sin embargo, son menos estables y en general se mueven más lentamente por lo que su uso como robots de servicio es aún muy limitado.

Recientemente han surgido los robots autónomos aéreos (drones), en particular aquellos basados en estructuras de tipo helicóptero como los cuadrópteros. Estos tienen también aplicaciones de servicio, como la de la entrega de paquetes a domicilio y la de vigilancia. En México hay algunos grupos que empiezan a explorar este campo, incluyendo grupos en el CINVESTAV, el ITESM y el INAOE.

4.7.2. Programas de apoyo

Los programas de apoyo proveen la interfaz entre los algoritmos de alto nivel y el hardware del robot, en particular los sensores y actuadores. Consisten de un conjunto de manejadores (*drivers*) y librerías que facilitan el acceso a cada tipo de dispositivo del robot y normalmente dependen de la plataforma robótica y dispositivo específico.

Adicionalmente, los programas de apoyo deben incluir los servicios análogos a los que proveen los sistemas operativos en las computadoras, como el manejo de la memoria, la ejecución y comunicación de procesos, la administración de usuarios, etc. Para ello se han utilizado sistemas operativos desarrollados para equipo de cómputo genérico, en particular sistemas basados en UNIX como LINUX orientados a los sistemas que funcionan en tiempo real. Recientemente han surgido sistemas operativos enfocados a robots, dentro de los cuales ROS (*Robot Operating System*)⁶⁸ se ha convertido en un estándar de facto en la comunidad robótica, en especial entre académicos. ROS facilita la interoperabilidad de los programas en diferentes plataformas y además per-

⁶⁸ <http://www.ros.org/>

mite compartir algoritmos entre los grupos de investigación, promoviendo la colaboración y el desarrollo de la robótica en el mundo.

4.8. Resumen y retos futuros

Los robots son máquinas que pueden percibir y actuar en el mundo controlados por procesos computacionales para realizar diversas tareas. Se pueden distinguir diferentes tipos de robots de acuerdo a su autonomía y la complejidad de su entorno, destacando los robots de servicio de media o alta autonomía que operan en entornos de complejidad baja o moderada. Estos robots se orientan a asistir a las personas en diversas actividades y ambientes, los cuales se pueden conceptualizar en tres niveles de sistema: i) funcional, ii) de dispositivos y algoritmos y iii) de implementación. El nivel funcional corresponde a la especificación y coordinación de las capacidades del robot; se han desarrollado diferentes esquemas a nivel funcional incluyendo las máquinas de estado, los lenguajes de especificación de tareas y los procesos de decisión de Markov. El nivel de algoritmos incluye los elementos computacionales que proveen las principales capacidades al robot: i) percepción y conducta motora, ii) audición y lenguaje, iii) conocimiento y razonamiento. El nivel de implementación está constituido por los elementos mecánicos y electrónicos del robot, y por los programas de apoyo y sistema operativo.

En México hay varios grupos activos en investigación y desarrollo en robótica, que han hecho aportaciones en los diferentes niveles y aspectos, con énfasis en los robots de servicio. Una forma de impulsar y evaluar los desarrollos en robótica es la participación en competencias, donde destaca RoboCup⁶⁹ a nivel internacional, en la cual participan regularmente los grupos mexicanos, y el Torneo Mexicano de Robótica⁷⁰ en nuestro país.

El desarrollo de los robots de servicio está en sus inicios y se espera que estos artefactos en el futuro sean ubicuos en nuestras casas, oficinas, hospita-

⁶⁹ <http://www.robocup.org/>

⁷⁰ <https://femexrobotica.org/>

les, etc., como los televisores y las computadoras. Sin embargo, para lograrlo todavía existen retos importantes por conquistar, incluyendo los siguientes:

- En cuanto a los aspectos mecánicos y electrónicos, se necesitan robots que puedan desenvolverse en ambientes como los que habitamos los humanos, para lo que se requiere robots humanoides que sean más robustos y rápidos que los actuales, con capacidades de desplazarse en ambientes complejos y manipular diferentes tipos de objetos. Deben mejorarse los sistemas de almacenamiento de energía (baterías) para que tengan una mayor autonomía.
- Aunque las capacidades de percepción, planeación y manipulación han evolucionado en los últimos años todavía falta mucho por hacer; por ejemplo que los robots puedan doblar prendas de ropa como una persona, abrir cajones y tomar objetos ocluidos, o desplazarse en ambientes desconocidos y dinámicos, entre otros.
- Dado que los robots de servicio tendrán que resolver tareas muy diversas, una capacidad fundamental es que puedan aprender de su experiencia o con ayuda del usuario, de forma que puedan realizar nuevas tareas o mejorar las que ya conocen. También deberán poder utilizar, como nosotros, los recursos disponibles en Internet, incluyendo textos, fotos y videos.
- La interacción con los humanos es fundamental para los robots de servicio, incluyendo la interacción a través del lenguaje natural hablado apoyada por ademanes. Además, para que logren una mayor empatía con sus usuarios deberán ser capaces de reconocer su estado emocional y simular ciertas emociones.